



# Text Detection

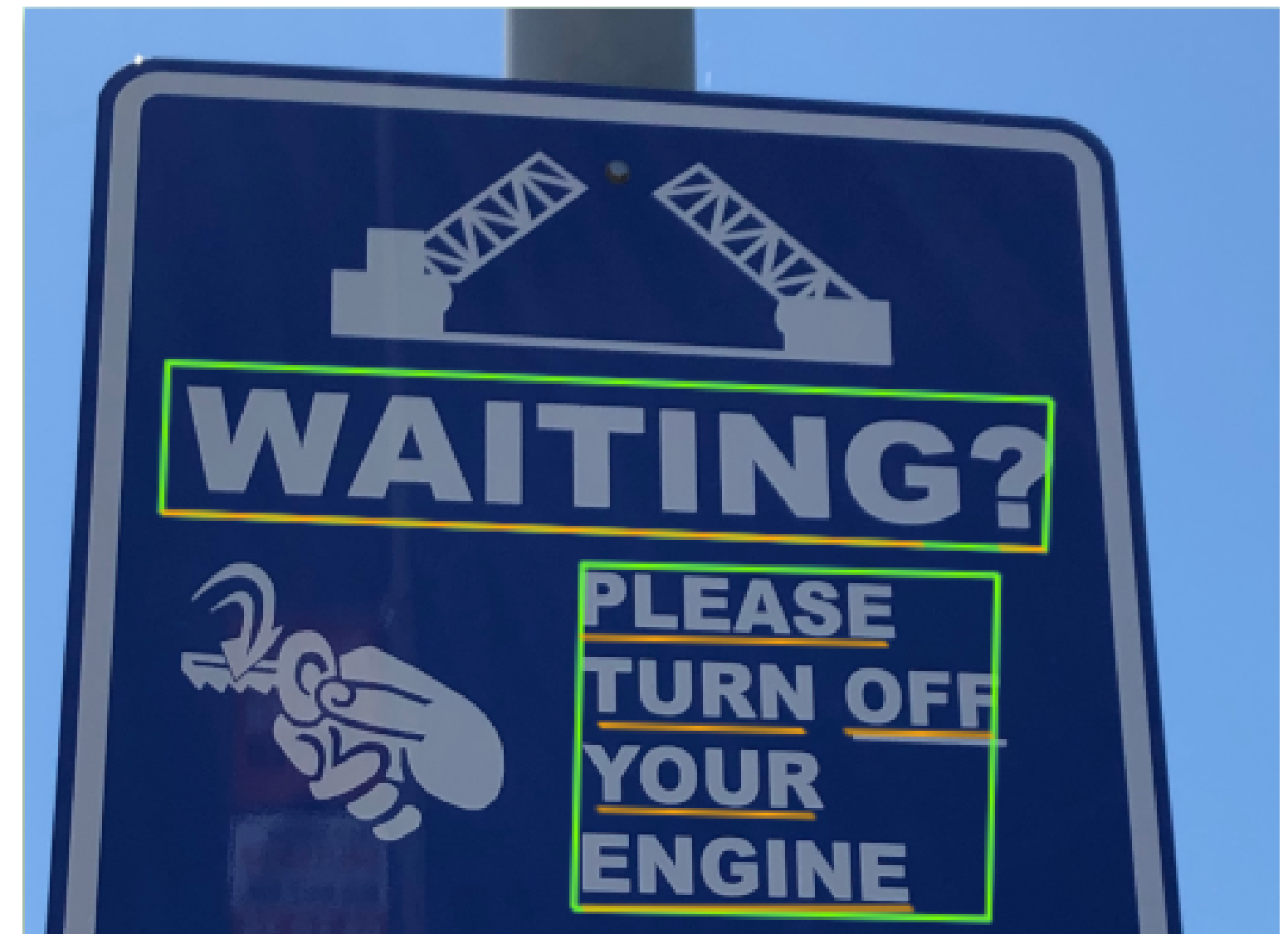
Jelvinlon Brilliando 2702231893

Muhammad Robby Kusumah 2702380891

# LATAR BELAKANG

## Text Detection

Kelompok kami memilih topik text detection karena teknologi ini merupakan solusi inovatif yang mengatasi berbagai permasalahan krusial di era digital saat ini. Text detection, atau teknologi pendeteksian teks, telah berkembang menjadi komponen penting dalam ekosistem teknologi modern sebagai respons terhadap beberapa tantangan signifikan yang dihadapi masyarakat global.



# Dataset

Kelompok kami menggunakan textOCR dataset yang berisi banyak gambar dengan text annotationnya

DATASETS

textocr-text-extraction-from-images

train\_val\_images

train\_images

0000599864fd15b3.jpg

0000e8b36676338b.jpg

00010bf498b64bab.jpg

00012db06b98f573.jpg

00016982f6086d39.jpg

0001a2f47097ee27.jpg

0001c6bf48e16ab2.jpg

000209d575f3aa4f.jpg

000228608388803f.jpg

0002c799b0cd7412.jpg

0002cb8d8ea5eb7e.jpg

0002d070329eb0fc.jpg

0002f9c7fac5f093.jpg

00035b31f27829d6.jpg

000384bb6da4764b.jpg

000421a4ed497ea4.jpg

00043e7c5c4c025a.jpg

00048f37069b6aa8.jpg

0004c9478eeda995.jpg

00054dab88635bdb.jpg

0005dcda3040dab3.jpg

00066af7b52f505c.jpg

00068a9d975dcc8a.jpg

0006dc0977056410.jpg

0007a5a18213563f.jpg

0007b54189a67423.jpg

0007f2c11800286e.jpg

00081dda0037f67.jpg

0008860dad1f9510.jpg

000957f9cbd14b35.jpg

25089 more

TextOCR\_0.1\_train.json

annot.csv

Input Data

TextOCR\_0.1\_train.json (279.89 MB)

About this file

Raw original json labels

This preview is truncated due to the large file size. The number of JSON items and individual items might be might be truncated. Create a Notebook or download this file to see the full content.

Download

▼ "root" : { 2 items

▶ "info" : {} 0 items

▼ "imgs" : { 747 items

▼ "a4ea732cd3d5948a" : { 5 items

"id" : string "a4ea732cd3d5948a"

"width" : int 840

"height" : int 1024

"set" : string "train"

"file\_name" : string "train/a4ea732cd3d5948a.jpg"

}

▶ "4bf43a7b2a898044" : {...} 5 items

▼ "1b55b309b0f50d02" : { 5 items

"id" : string "1b55b309b0f50d02"

"width" : int 1024

"height" : int 683

"set" : string "train"

"file\_name" : string "train/1b55b309b0f50d02.jpg"

}

▶ "00c359f294f7dcd9" : {...} 5 items

▼ "04b5a37f762b0f51" : { 5 items

"id" : string "04b5a37f762b0f51"

"width" : int 768

"height" : int 1024

"set" : string "train"

"file\_name" : string "train/04b5a37f762b0f51.jpg"

}

▶ "b3c4d860c380c674" : {...} 5 items

▶ "edecba688d135d36" : {...} 5 items

▶ "01ee1b7395e1d8c4" : {...} 5 items

▶ "344193868ca2f738" : {...} 5 items

▶ "a183473b9b4bc8f6" : {...} 5 items

▶ "4e40fcd84677d90" : {...} 5 items

▶ "c126876b5ac73c91" : {...} 5 items

▶ "57ad2d3ddcde298b" : {...} 5 items

▶ "171e532e7edaf8f9" : {...} 5 items

▶ "754dd4b30ec02ca5" : {...} 5 items

▶ "a0fa95532b52cf18" : {...} 5 items

▶ "d3abbec0a3c582c1" : {...} 5 items

▶ "75cd66309f4183ae" : {...} 5 items

Input Data

000209d575f3aa4f.jpg (568.72 kB)

About this file

training image



Input (7.58 GB)

Data Sources

TextOCR - Text Extraction

train\_val\_images

TextOCR\_0.1\_train.json

annot.csv

annot.parquet

img.csv

img.parquet



# Contoh Gambar

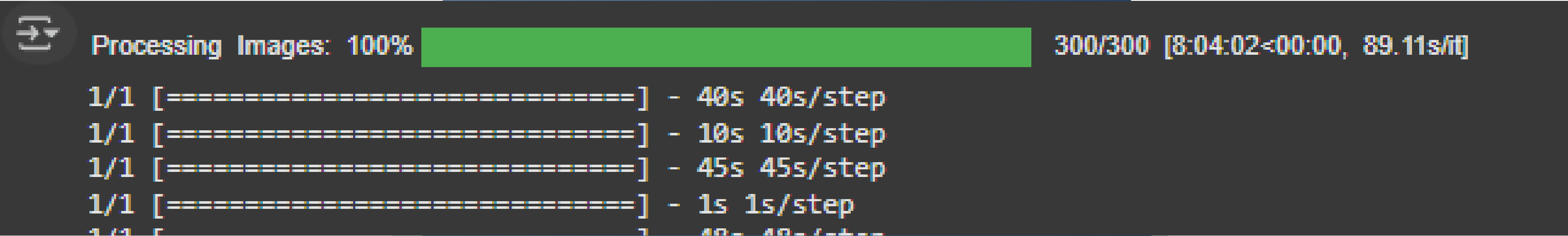
```
[8]: fig, ax = plt.subplots(figsize=(10, 10))  
     ax.imshow(plt.imread(img_fns[20]))  
     ax.axis('off')  
     plt.show()
```



+ Code

+ Markdown

# Hasil menjalankan 100 set gambar



--- Results Preview ---

	file_name	ground_truth	easyocr_text	easyocr_accuracy	keras_ocr_text	keras_ocr_accuracy	pytesseract_text	pytesseract_accuracy
0	3a7d196f293b0bf5.jpg	4-9% HATLIFTER STOUT GRAND RIDGE BREWERY VINTAGED NO ADDED CHEMICALS OR PRES...	(ewer ~u5 Most AWARDED 'BREWERY 49" euryA GRAND 9 9 .ko Rewe?) 6 Aunt Aobrd...	0.361775	crans ridge srewer brewery o most awarded wolift g crand ridge ppewev wo tre...	0.412698		0.000000
1	1940f68d6d697475.jpg	Ping Pin LEMON . -So So Pin . Pi S P . . . . .	LEd Pin PvSe Ord ES	0.338462	dv lemol so p s din	0.400000		0.000000
2	73db6439b8b9aa38.jpg	adidas . . das .	0S odlaas	0.400000	css caigas	0.307692		0.000000
3	01a28309568b4274.jpg	Chan ans 19 TOEWS KANE 88 88	TOEwS 19 8 KINE Ohay	0.408163	toews 19 o kane cha ans	0.509804		0.000000
4	3b3fb77c202ef6eb.jpg	2nd Edition COMPUTER ALGORITHMS HOROWITZ SAHNI RAJASEKERAN SP AN INTRODUCTIO...	A LGOR T5 COMPUTER Edition ALGORITHMS COvPUTERS AND InTRACTABILITY Guda Kotn...	0.481481	algorts 2nn edition computer computers and intractability guce witing lo ino...	0.466312	CompurTerR\nALGORITHMS\n\nHorowmz + Sauni\nRAJASEKERAN\n\nTrogramming\n\n%\n\n...	0.324519

--- Average Model Accuracy ---

EasyOCR Average Accuracy: 45.47%

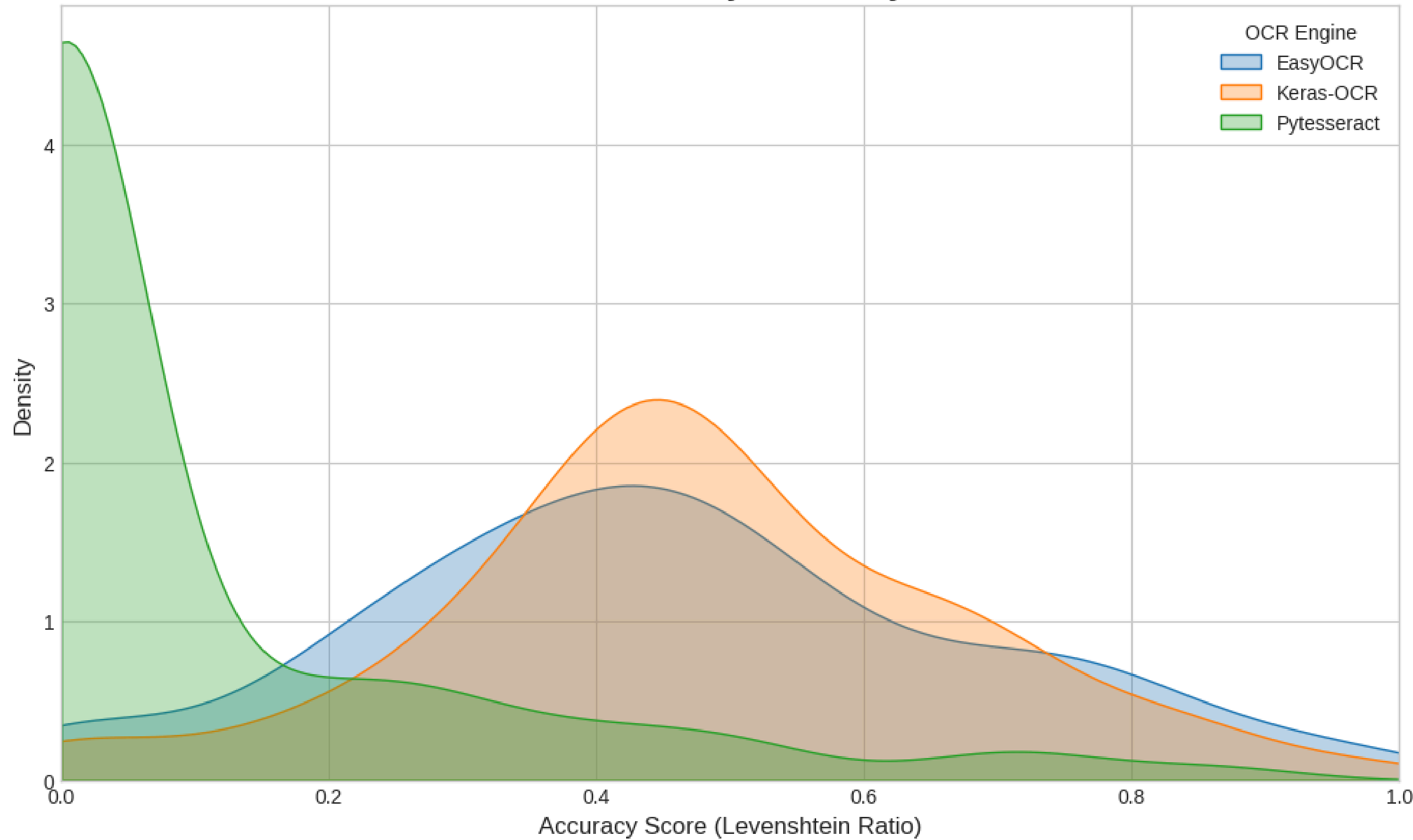
Keras-OCR Average Accuracy: 47.56%

Pytesseract Average Accuracy: 10.10%

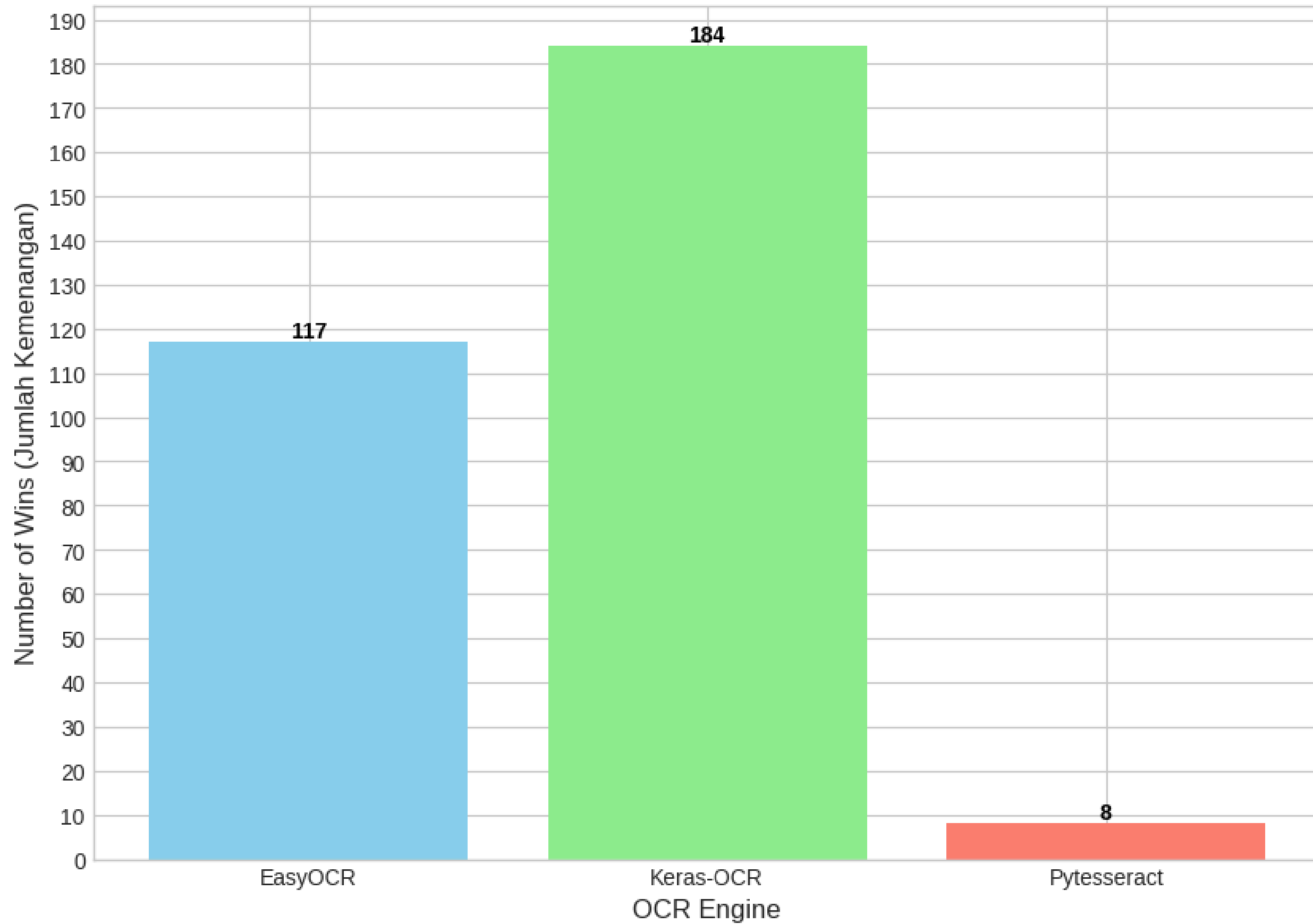
Full detailed results have been saved to 'ocr\_comparison\_results.csv'

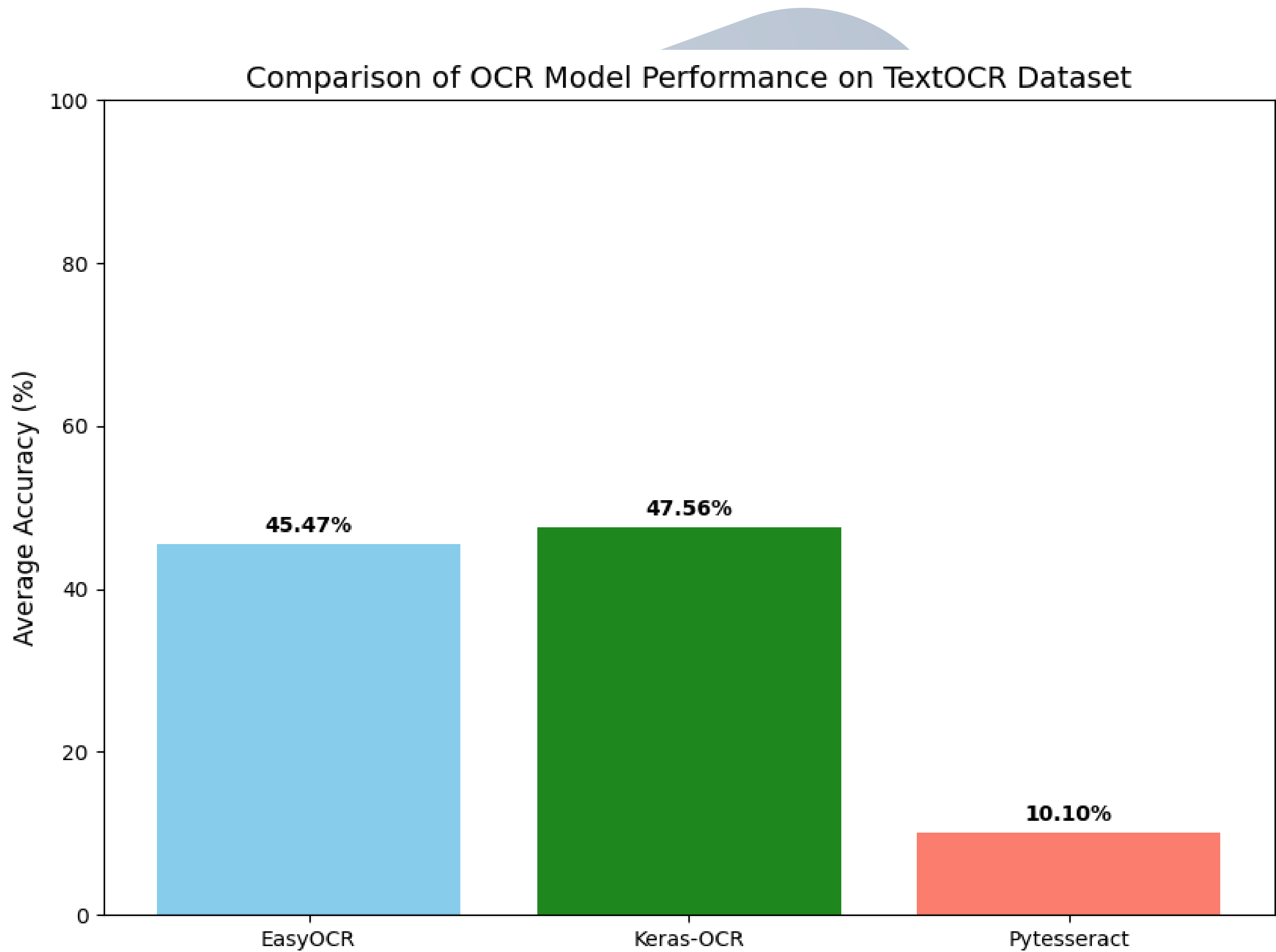
You can download this file from the file browser on the left sidebar.

Distribution of Accuracy Scores by OCR Model



Number of Times Each OCR Engine Achieved Highest Accuracy







# Implementasi ResNet dalam EasyOCR

EasyOCR menggunakan ResNet sebagai backbone untuk komponen recognizer

Biasanya menggunakan ResNet-50 yang telah dimodifikasi

Output feature maps dari ResNet diproses oleh:

Bidirectional LSTM untuk sequence modeling

CTC (Connectionist Temporal Classification) untuk decoding

Model dilatih pada dataset besar yang berisi berbagai jenis teks

# Keunggulan ResNet dalam Deteksi Teks

Ketahanan terhadap variasi visual (robust feature extraction)  
Kemampuan generalisasi yang baik untuk berbagai jenis teks  
Pembelajaran hierarkis:

Layer awal: garis, tepi, kurva

Layer tengah: komponen karakter

Layer akhir: representasi karakter lengkap